Optimization of Hate Speech and Abusive Language Detection on Indonesian-language Twitter using Genetic Algorithms

Authors : Rikson Gultom

Abstract : Hate Speech and Abusive language on social media is difficult to detect, usually, it is detected after it becomes viral in cyberspace, of course, it is too late for prevention. An early detection system that has a fairly good accuracy is needed so that it can reduce conflicts that occur in society caused by postings on social media that attack individuals, groups, and governments in Indonesia. The purpose of this study is to find an early detection model on Twitter social media using machine learning that has high accuracy from several machine learning methods studied. In this study, the support vector machine (SVM), Naïve Bayes (NB), and Random Forest Decision Tree (RFDT) methods were compared with the Support Vector machine with genetic algorithm (SVM-GA), Nave Bayes with genetic algorithm (NB-GA), and Random Forest Decision Tree with Genetic Algorithm (RFDT-GA). The study produced a comparison table for the accuracy of the hate speech and abusive language detection model, and presented it in the form of a graph of the accuracy of the six algorithms developed based on the Indonesian-language Twitter dataset, and concluded the best model with the highest accuracy.